THESIS

RELATING SEVERITY OF A MOUNTAIN PINE BEETLE OUTBREAK TO FOREST MANAGEMENT HISTORY

Submitted by

Anthony Vorster

Graduate Degree Program in Ecology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2014

Master's Committee:

Advisor: Thomas Stohlgren
Co-Advisor: Sunil Kumar

Tony Cheng
Paul Evangelista

ABSTRACT

RELATING SEVERITY OF A MOUNTAIN PINE BEETLE OUTBREAK TO FOREST MANAGEMENT HISTORY

The availability of remote sensing imagery before, during, and after the recent mountain pine beetle (*Dendroctonus ponderosae* Hopkins) epidemic in the southern Rocky Mountains presents exciting opportunities for assessing the current state of forests and how forest management in previous decades influenced outbreak severity across the landscape. I mapped outbreak severity at a 30-m resolution using integrative spatial modeling. I predicted that: 1) outbreak severity can be accurately predicted and mapped at Fraser Experimental Forest, Colorado using stand characteristics with a boosted regression trees model, Landsat imagery, geographic information system (GIS) data, and field data; and 2) forest stands that were unmanaged since the 1950s will have higher outbreak severity compared to stands that were treated since the 1950s. Outbreak severity, measured by the ratio of dead lodgepole pine (*Pinus contorta*) basal area to the basal area of all trees, was mapped across Fraser Experimental Forest with a cross-validation correlation of 0.86 and a Spearman correlation with independently observed values of 0.64. The outbreak severity at stands harvested between 1954 and 1985 was lower than comparable uncut stands. Lessons learned about past treatments will inform forest management for future mountain pine beetle outbreaks.

ACKNOWLEDGEMENTS

I would like to thank everyone involved in the development of this research. Renee Curry, Kelli Groy, Aaron Sidder, and William Zawacki contributed to the development of the models. Steve Chignell, Matthew Luizza, Tewodros Wakie, and Amanda West helped collect field data. Wayne Shepperd shared knowledge of forest management history at Fraser Experimental Forest. Robert Hubbard, Charles Rhoades, and Kelly Elder at the USDA Forest Service Rocky Mountain Research Station provided field data and valuable background information. My research was strengthened by conversations with Nicholas Young and John Hogland. Lastly, this work would not be possible without the support and feedback from my committee members: Tony Cheng, Paul Evangelista, Sunil Kumar, and Thomas Stohlgren. Funding was provided by the Bioenergy Alliance Network of the Rockies (BANR).

TABLE OF CONTENTS

LIST OF KEYWORDS

boosted regression trees, Landsat, mountain pine beetle, remote sensing, silviculture

1. INTRODUCTION

The mountain pine beetle (*Dendroctonus ponderosae* Hopkins) is a bark beetle native to western North America that feeds on and reproduces inside of its pine tree hosts. Lodgepole pine (*Pinus contorta)* serves as the primary host (Wulder et al., 2006a). Mountain pine beetles bore into a host tree and emit an aggregation pheromone that attracts other beetles in the area (Six et al., 2014). Trees employ a range of defenses against the beetles. However, when trees are vulnerable or there is a high population density of mountain pine beetles, beetles can overcome a tree's defenses (Raffa et al., 2008). They initially attack in August and lay eggs soon after. Larvae hatch and develop in the host tree and emerge the following summer. Host tree coloration indicates the time since attack. Trees do not show signs of attack initially during the period known as the green phase. Pine needles of infected trees gradually change from green to yellow to red for the year following the attack. The red phase is marked by the needles on the host tree turning red, typically one year after attack. Trees enter the gray phase as the red needles drop from the tree, exposing the bare branches (Wulder et al., 2006a). Mountain pine beetle behavior is directly related to its population size which is mediated by a series of climatic and biotic thresholds (Raffa et al., 2008). Mountain pine beetles at an endemic population size select weakened trees as hosts. Warm and dry weather along with mild winters provide favorable conditions for increased beetle population size (Safranyik, 2003). Increased population sizes allow for the beetles to overwhelm the defenses of even healthy trees. Outbreaks can be slowed by cold weather in the late fall to early spring or depletion of suitable host trees (Safranyik, 2003).

The recent mountain pine beetle outbreak has impacted approximately 1.4 million hectares of forests in Colorado since it began in 1996, but is now on the decline (Colorado State Forest Service, 2013). The outbreak of mountain pine beetle in western North America was estimated to be larger and more severe than historical outbreaks (Six et al., 2014). Anthropogenic factors such as increased forest homogeneity and changing climate have likely contributed to the increased magnitude (Raffa et al., 2008).

Forest managers have attempted to control the spread of the recent mountain pine beetle outbreak through a variety of methods. Direct control treatments aim to reduce or eradicate mountain pine beetle populations, while indirect control methods attempt to reduce beetle populations by promoting improved tree vigor and stand conditions unfavorable for beetle populations (Six et al., 2014). Once mountain pine beetles reach epidemic-level population sizes, these control methods are generally ineffective at the scales and intensity they are currently implemented (Six et al., 2014). Safranyik (2003) recommended that management efforts should focus on lodgepole pine rather than on mountain pine beetle. Harvesting of lodgepole pine stands is one strategy that has been implemented to indirectly manage mountain pine beetles (Whitehead et al., 2003). I created a map of outbreak severity at Fraser Experimental Forest, Colorado, and analyzed the outbreak intensity in relation to characteristics of forest stands that were harvested in the past. Understanding the fate of historically managed stands informs current management. The management decisions of today will shape susceptibility of our forests to future outbreaks.

The severity of the mountain pine beetle outbreak at a given location has implications for forest regeneration (Diskin et al., 2011), hydrology (Raffa et al., 2008), fire (Simard et al., 2011), timber harvest (White et al., 2005), and carbon accounting (Pfeifer et al., 2011). The United

States Department of Agriculture (USDA) Forest Service Aerial Detection Surveys show where and when mountain pine beetle outbreaks occurred (USDA Forest Service 2013). These surveys are useful at coarse management and study scales; however, finer resolution maps showing the gradient of severity across the landscape are needed to set management priorities for post-infestation Rocky Mountain forests (Pfeifer et al., 2011). My first objective was to map mountain pine beetle outbreak severity as measured by the ratio of dead lodgepole pine basal area to basal area of all trees in a 30 x 30 m pixel. A model that achieved a cross-validation correlation of 0.70 and a correlation of 0.70 with observed values from independent field data would be considered accurate. Such a map would have many potential applications; I specifically used this map to analyze how characteristics of previously harvested stands correlated to outbreak severity.

Remote sensing has been frequently used to study the recent mountain pine beetle outbreak (Wulder et al., 2006a). Previous work at the landscape scale has mapped extent and timing of the outbreak (Goodwin et al., 2008; Coops et al., 2010; Walter and Platt, 2013), probability of red attack (Wulder et al., 2006b), and red attack damage categories (Skakun et al., 2003). Landsat TM and ETM+ have been employed to consistently map red attack with 70-75% accuracy (Wulder et al., 2006a). It is suggested that the goal of landscape scale mountain pine beetle mapping efforts should be a low-cost option to mapping severity of mortality (Wulder et al., 2006a). Outbreak severity has been mapped with high-resolution imagery that may not necessarily be practical at landscape scales (Dennison et al., 2010). Recent work used a time series of Landsat imageries to map outbreak severity and study outbreak dynamics (Meddens and Hicke, 2014). My work continues this effort to map outbreak severity using Landsat ETM+ imagery. It is novel in its mapping of a continuous spectrum of mountain pine beetle outbreak severity for each pixel using Landsat ETM+ imagery to detect changing forest conditions, not

red phase trees. My work uses freely-available satellite imagery, geospatial data, and geographic information system (GIS) software. Anyone with access to field data and geospatial training can reproduce these methods to determine outbreak severity in their area of interest. Integrative spatial models use a variety of data types to create powerful, inexpensive, and spatially explicit results.

Silviculture prescriptions can reduce the chances of tree mortality during a mountain pine beetle epidemic (Whitehead et al., 2003). Stand conditions associated with high mountain pine beetle induced mortality are those that attract mountain pine beetles and reduce tree vigor (Hicke and Jenkins, 2008). In general, susceptible stands have large lodgepole pine trees and/or are dense. Harvests can reduce both the average size and density of host trees. The literature suggested that lodgepole pine stands <60 years old are not impacted by outbreaks, stands between 60-80 years old are rarely infested, and stands >80 years old are commonly highly susceptible to mountain pine beetle attack (Shore and Safranyik, 1992). Following this logic, my second objective was to investigate if forest stands that were actively managed in the decades leading up to the recent mountain pine beetle epidemic had lower mortality than unmanaged, older areas. Additionally, I asked what characteristics of harvested stands influenced the levels of outbreak severity. I hypothesized that:

1) Outbreak severity can be accurately predicted at Fraser Experimental Forest using pre, mid, and post-outbreak stand characteristics with a boosted regression trees (BRT) model, Landsat imagery, GIS data, and field data.
2) Forests managed since the 1950s will have lower levels of mortality than stands in neighboring watersheds.
3) The more recently harvested the forest, the lower the mortality level.

4) Time since harvest will be the most important variable in explaining outbreak severity in cut stands.

2. MATERIALS AND METHODS

2.1. Study Area

Fraser Experimental Forest, Colorado is managed by the USDA Forest Service and has an active research history. Elevation ranges from 2,650 m to 3,900 m and temperature varies along the elevation gradient from -40°C to 32°C (Collins et al., 2011). The average precipitation at Fraser Experimental Forest is 71-76 cm; two thirds of this precipitation falls as snow between October and May. The vegetation is typical of the central Rocky Mountains. Lower elevations and southern aspects are dominated by lodgepole pine forests. Mixed forests of Engelmann spruce (*Picea engelmannii*), sub-alpine fir (*Abies lasiocarpa*), and lodgepole pine are found along the valley bottom and north-facing slopes up to treeline (3300-3500 m; Collins et al., 2011). Small stands of quaking aspen (*Populus tremuloides*) are interspersed throughout lower elevations. Mountain pine beetles started to cause widespread lodgepole pine mortality in 2002 in Fraser Experimental Forest (Collins et al., 2011**)**.

2.2 Delineation of Past Harvests

The timber harvests used in my work were experimental harvests cut for various studies between 1954 and 1985 (Alexander et al., 1985; Troendle and King, 1985). This range of harvest years was selected because of the availability of information and because stands cut between 30 and 60 years ago are the stands that would not be expected to be impacted by mountain pine beetle (Shore and Safranyik, 1992). Two areas with high concentrations of cuts are part of paired watershed studies. The cuts in Fool Creek watershed were a series of clearcuts of varying widths covering 113 hectares cut between 1954 and 1956. Circular clearcuts on the Deadhorse Creek

watershed were cut between 1977 and 1978. Additional irregular-shaped clearcuts were completed on the Deadhorse Creek watershed between 1982 and 1984 (Alexander et al., 1985).

Details about past harvests were pieced together from a technical report (Alexander et al., 1985), aerial photos shared by the USDA Forest Service Rocky Mountain Research Station, United States Geological Survey (USGS) National Aerial Photography Program (NAPP) photos, USDA Farm Service Agency National Agriculture Imagery Program (NAIP) photos, and personal communication with Dr. Wayne Shepperd (personal communication, May 15, 2014). These sources informed the delineation of past harvests in ENVI (Exelis Visual Information Solutions, 2014) image analysis software. ArcGIS (version 10.1; ESRI, 2012) software and the sources mentioned above were then utilized to compile information about the age, size, slope, aspect, and elevation of each harvest. Three harvests just outside of Fraser Experimental Forest were within the extent of the imagery clipped to the rectangular extent of Fraser Experimental Forest and were included because they provided data for cuts larger than most cuts at Fraser Experimental Forest. Estimates of the age of each cut were as accurate as possible given the available information. The age of most cuts was accurate within one or two years, but there are a few harvested stands where the age is accurate to seven years. Treatments that were only 20 m wide in Fool Creek watershed and other small treatments in other watersheds were not delineated because geometric registration and resolution of Landsat combined with potential harvest delineation errors would compromise the accuracy of analysis. The harvest polygons I created represent the majority of cuts completed between 1954 and 1985, but do not include every cut completed during this timeframe.

2.3 Field Data

The USDA Forest Service surveyed 74 17.9 m fixed-radius plots in the summer of 2008. Plot locations were generated randomly in the forested areas of Fraser Experimental Forest. Species, diameter at breast height (dbh), status (canopy class or dead), and evidence of mountain pine beetle were recorded for each tree >5 cm dbh. Of the 73 plots, 48 of them located in lodgepole pine forests were revisited in the summers of 2012 and 2013. The 48 plots sampled in 2012 and 2013 were used with 25 plots from 2008 to train the outbreak severity BRT model. These data will be referred to as the training data. Although the outbreak was still active between 2008 and 2012 (USDA Forest Service, 2013), the 25 plots from 2008 were included because they were either located in subalpine fir and Engelmann spruce stands that are not susceptible to mountain pine beetle or already had very high levels of lodgepole pine mortality by 2008.

A second field dataset was used as independent test data to validate the outbreak severity BRT model. Snow courses were established in four watersheds at Fraser Experimental Forest as part of a paired watershed study to examine the influence of timber harvest on snow accumulation, sediment production, and yield and timing of streamflow (Troendle and King, 1985). A total of 133 plots across the four watersheds were surveyed in 2007 to record species, dbh, status (crown class or dead), and year of mountain pine beetle caused mortality for each tree in 10 m fixed-radius plots. These plots were revisited in 2011 to record any changes in tree status. There are an average of 33 plots organized in a looping pattern across all aspects, elevations, and stand conditions in each watershed (Troendle and King, 1985). The plots are roughly evenly spaced an average of 141 m apart in Deadhorse Creek, 159 m apart in East Saint Louis Creek, 162 m apart in Fool Creek, and 83 m apart in Lexen Creek.

The ratio of dead lodgepole pine basal area to the basal area of all trees, alive and dead, was calculated at each training and test plot as a measurement of mountain pine beetle outbreak severity. Basal area was calculated for each tree using Equation 1 and then the basal area of dead lodgepole pine and of all trees was summed by plot.

$$BA=0.00007854*dbh^2 \quad (1)$$ (Walker et al., 2007)

Where:

BA= basal area in square meters

dbh= diameter at breast height in centimeters

The difference in plot area between the training (17.9-m radius) and test (10-m radius) was accounted for because the response variable was a ratio and the plot area cancelled out of the numerator and denominator of the ratio. Using the ratio of dead lodgepole pine basal area to the basal area of all trees as a measurement of mountain pine beetle outbreak severity assumes that all dead lodgepole pine was a result of mountain pine beetles. Observations from the field indicate that the majority of dead lodgepole pine in these plots is a result of the mountain pine beetle attack.

2.4 GIS and Remote Sensing Data

GIS and remote sensing data were used as predictor variables for the mountain pine beetle outbreak severity BRT model. Specifically, elevation, slope, and aspect were created from a 10-m resolution digital elevation model (DEM) from the USGS National Elevation Dataset (Gesch et al., 2002; Gesch, 2007). I converted aspect from a circular scale (degrees) to a linear north-south (northness) and east-west (eastness) gradient (Kumar et al., 2006). This was done using a cosine and sine transformation for northness and eastness, respectively (Morrison et al.,

2003). Once converted, aspect was defined by two values ranging from -1 (south or west) to 1 (north or east). Topographic variables are important predictors when detecting mountain pine beetle-induced mortality because topography represents climatic conditions, such as moisture and temperature, which influence beetle outbreaks (Nelson et al., 2013).

Landsat imagery is frequently used to map mountain pine beetle activity because of its moderate 30-m spatial resolution, spectral resolution, and the ability to use historic Landsat images to document past vegetation conditions (Goodwin et al., 2008). The 30-m resolution of Landsat Enhanced Thematic Mapper Plus (ETM+) and the spectral resolution are suitable for detecting patterns from mountain pine beetle at the epidemic phase (Bentz and Endreson, 2003). Landsat ETM+ images from September were collected because cloud-free images were available before (2001), during (2010), and after (2012) the height of the outbreak. All images were collected from the same month so the phenology of the landscape was about the same (Lu et al., 2004). Scenes collected from different months may lead to misclassification of areas that are changing due to seasonality rather than lodgepole pine mortality. Additionally, September was chosen because it allows time for the mountain pine beetle host trees to show up as dead from the previous year's attack (Goodwin et al., 2008). Landsat 7 ETM+ images were downloaded from the USGS Earth Explorer website for September 24, 2001, September 17, 2010, and September 22, 2012. A mask was created to remove a small area in the 2012 image that is blocked by cloud cover. Fraser Experimental Forest lies within the unaffected portion of the Landsat 7 ETM+ images after the Scan Line Corrector (SLC) failed in May 2003.

Certain steps should be taken before change detection analysis can be implemented to ensure changes observed in the imagery truly reflect changes in the landscape: geometric registration, radiometric and atmospheric calibration, and topographic correction if working in

mountainous areas (Lu et al., 2004). Level 1T processed Landsat images have already undergone geometric registration and topographic correction. Radiometric calibration was performed in ENVI to convert digital numbers to radiance. Atmospheric scattering was then accounted for using dark object subtraction (Song et al., 2001). Images were clipped to the extent of Fraser Experimental Forest.

The processed Landsat ETM+ images were used to generate continuous surfaces of predictor variables for the BRT model. I created the predictor variables used in both the BRT model predicting outbreak severity and the regression tree model exploring the relationship between harvested stand characteristics and outbreak severity (Table 1). All image processing, calculation of indices and textures, and image differencing was done in ENVI.

Table 1. Summary of response variables, predictor variables, and models used.

Objective	Response Variable	Source of Response	Predictor Variables	Model
Map Outbreak Severity	Dead Lodgepole Pine Basal Area/Total Basal Area	Training Plots	Aspect Elevation Landsat Bands NDMI Difference Tasseled Cap Difference Texture Difference Slope	Boosted Regression Trees (BRT)
Historical Management Analysis	Average Dead Lodgepole Pine Basal Area/Total Basal Area of Pixels Within Harvest Polygon	Outbreak Severity Map Produced by BRT	Aspect Elevation Slope Size of Treatment Years Since Cut	Regression Tree

The Normalized Difference Moisture Index (NDMI) and its difference over time was selected as a predictor variable because it has been used to effectively identify mountain pine beetle-killed stands (Goodwin et al., 2008; Coops et al., 2010; Walter and Platt, 2013). NDMI was calculated using the following equation:

$$NDMI=(NIR-MIR)/(NIR+MIR) \quad (2)$$ (Goodwin et al., 2008)

Where:

NDMI= Normalized Difference Moisture Index

NIR= Near-Infrared, ETM+ Band 4

MIR= Mid-Infrared, ETM+ Band 5

The Tasseled Cap Transformation has been used to map forest attributes and the change in Tasseled Cap values has been used to map mountain pine beetle outbreaks (Skakun et al., 2003; Wulder et al., 2006b). Three Tasseled Cap bands were calculated: brightness, greenness, and wetness (Crist and Cicone, 1984; Huang et al., 2002) using the following equations:

Brightness=0.3651(ETM+1)+0.3972(ETM+2)+ (3) (Huang et al., 2002)

0.3904(ETM+3)+0.6966(ETM+4)+0.2286(ETM+5)+0.1596(ETM+7)

Greenness=(-0.3344)(ETM+1)+(-0.3544)(ETM+2)+ (4) (Huang et al., 2002)

(-0.4556)(ETM+3)+0.6966(ETM+4)+(-0.0242)(ETM+5)+(-0.2630)(ETM+7)

Wetness=0.2626(ETM+1)+0.2141(ETM+2)+ (5) (Huang et al., 2002)

0.0926(ETM+3)+0.0656(ETM+4)+(-0.7629)(ETM+5)+(-0.5388)(ETM+7)

Where:

ETM+=Enhanced Thematic Mapper Plus and the number indicates the band number

Image texture for ETM+ bands 3, 5, and 7 can provide information about forest stands and biomass (Lu and Batistella, 2005; Singh et al., 2014). Texture was calculated using the Co-occurrence Measures Tool in ENVI. I calculated texture for bands 3, 4, and 5 since these bands were likely to contain information about vegetation and moisture that could reveal patterns in mountain pine beetle outbreak severity. Mean and homogeneity textures for bands 3, 4, and 5 were calculated using three window sizes: 3x3, 5x5, and 7x7.

Stands that experienced high levels of mortality during the mountain pine beetle outbreak are expected to change more than stands with low levels of lodgepole pine mortality. Image

differencing was used to highlight these areas of change. All Landsat ETM+ bands, NDMI, Tasseled Cap Bands, and textures were differenced from 2001 to 2010, 2001 to 2012, and 2010 to 2012. These time steps were used in attempt to capture changes in stand conditions across the time frame of interest and also in smaller increments throughout the outbreak.

Values from each predictor variable raster image were extracted at each training and test plot location using the USGS VisTrails Software for Assisted Habitat Modeling (SAHM; Morisette et al., 2013). SAHM was then used to test correlations between predictor variables. The large number of predictor variables was narrowed down using an iterative process of testing various combinations of predictor variables in the BRT model, ecological interpretation and checking model performance, variable importance, and predictor variable correlations. If predictor variables were correlated by 0.70 or more, the variable with the most predictive power and that made the most ecological sense was kept and the other predictor was discarded. I used no more than seven predictor variables so the number of predictor variables did not exceed one tenth the number of training plots (Harrell et al., 1996).

2.5 Forest Mask

Model predictions of mountain pine beetle outbreak severity are only valid across the forested areas of Fraser Experimental Forest. I created a mask of forested areas using a Maximum Likelihood supervised classification in ENVI. The classification was trained with polygons derived from visual inspection of NAIP imagery.

2.6 Statistical Analyses

For my first objective, I used a combination of field data, remote sensing, and GIS data to model mountain pine beetle outbreak severity. A BRT model was used to find the correlation between the response variable, the ratio of dead lodgepole pine basal area to basal area of all trees, and the predictor variables (Table 1). BRT use regression trees and boosting to create a model that uses different types of predictor variables, handles complex interactions, and predicts well (Elith et al., 2008). Regression trees model relationships between response and predictor variables using repeated binary splits in the data that reduce the amount of variation in each group. Boosting builds many regression trees in an ordered fashion that increases emphasis on poorly modeled observations (Elith et al., 2008). Code was modified from Elith et al. (2008) to implement the BRT in R statistical software using the Generalized Boosted Regression Models (GBM) package (Appendix; R Development Core Team, 2013). A learning rate of 0.0025, tree complexity of 3, and a bag fraction of 0.5 yielded a model of 1,150 trees. These parameter values were selected because they maximized model performance and were within the suggested range (Elith et al., 2008).

The predictions from the BRT were applied to the predictor variable rasters to create a continuous map of predicted outbreak severity which informed the analysis in my second objective. The value in each pixel represents the ratio of dead lodgepole pine basal area to the total basal area. The values of each pixel in a cut were averaged to get the average outbreak severity in each cut. The relationship between outbreak severity and characteristics of past treatments and topography were analyzed using a regression tree (Table 1). Regression trees were chosen for their ease of use and interpretation and their ability to handle a variety of data

types (De'ath and Fabricius, 2000). The regression tree was implemented in R statistical software using the Classification and Regression Trees (Tree) package.

3. RESULTS

3.1 Objective 1: Model Outbreak Severity at Fraser Experimental Forest

Models of outbreak severity performed well when tested with a cross-validation correlation, but did not perform as well when tested with the independent test data. The BRT model had a cross-validation correlation of 0.86. The predicted outbreak severity and observed outbreak severity at the test plots had a Spearman's correlation of 0.64 and an adjusted R^2 value of 0.43, indicating that 43 percent of the outbreak severity variation at the test plots can be explained by the BRT model.

The top two predictor variables of outbreak severity contained far more predictive power than the other predictors (Table 2). The model was simplified to five predictor variables because increasing the number of predictor variables complicated the model without increasing predictive power. The best predictors of outbreak severity involved image differencing between 2001 and 2010 and conditions in 2010, but no images from 2012 were used.

Table 2. The relative percent contribution of each predictor variable to the BRT model of outbreak severity.

Variable	Percent Contribution
ETM +Band 5 Mean Texture Difference (2001-2010) 3 by 3 Window	46
NDMI Difference (2001-2010)	36
ETM+ Band 5 (2010)	8
Tasseled Cap Brightness Difference (2001-2010)	5
Elevation	5

The resulting map of outbreak severity shows a range in the ratio of dead lodgepole pine basal area to total basal area at Fraser Experimental Forest from 0 (lowest) to 0.65 (highest) (Figure 1). The maximum predicted value of 0.65 is lower than the outbreak severity calculated at eight of the training data plots and 30 test data plots. The BRT model underestimated outbreak severity. A few trends stand out upon visual examination of the map. The yellow, orange, and red areas indicating higher levels of outbreak severity are concentrated at the lower elevations of Fraser Experimental Forest. A zone of green buffers the alpine zone and areas of high mortality indicating the presence of higher-elevation forests that had lower levels of outbreak severity. These patterns correspond to the lower elevation lodgepole pine forests and the higher elevation spruce and fir forests. Classification of the continuous outbreak severity map into bins of low, moderate and high mortality allows for quantification of forested area impacted by different levels of outbreak severity (Table 3). The majority of the forest experienced low or high levels of mortality, with only 12 percent of the forest experiencing moderate levels of lodgepole pine mortality (Table 3).

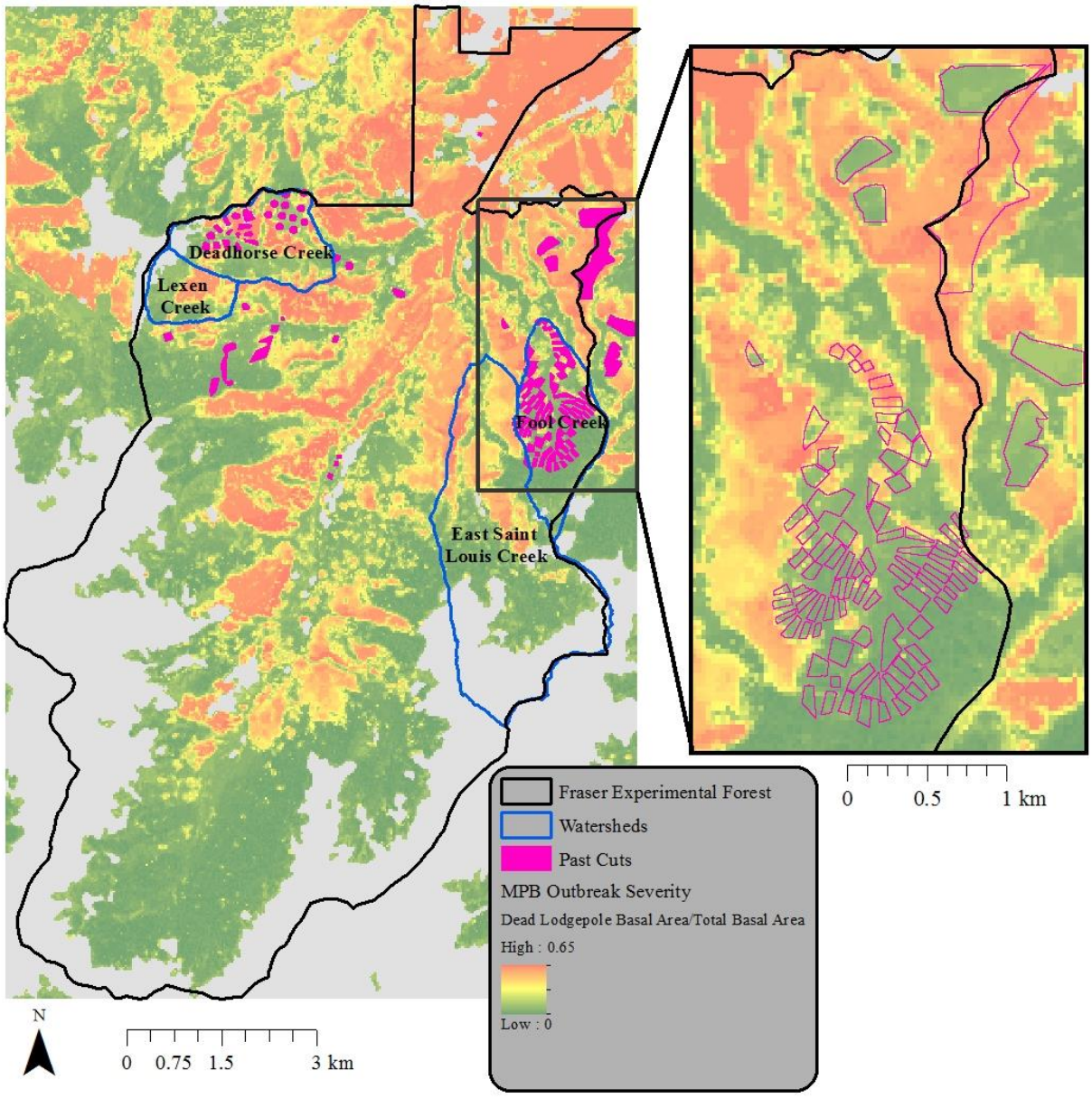

Figure 1. Predicted outbreak severity by the BRT model across Fraser Experimental Forest. The grey areas on the map were masked from the results because they are not forested. The map subset on the right shows the cuts in and around the Fool Creek Watershed. Notice that the boundaries of some of the cuts are clearly visible because they are pockets of low mortality surrounded by areas of high lodgepole pine mortality.

Table 3. Predicted area in different outbreak severity categories in Fraser Experimental Forest.

Classification	Dead Lodgepole Pine Basal Area/ Total Basal Area	Area (hectares)	Percent of Forested Area
Low Severity	0-0.20	3,966	57
Moderate Severity	0.21-0.40	846	12
High Severity	0.41-0.65	2,096	30

3.2 Objective 2: Analyze Relationship between Outbreak Severity and Historical Management

A comparison of outbreak severity between harvests in the Fool Creek and Deadhorse Creek watersheds and the entire drainages and control watersheds showed that harvested areas consistently had lower levels of outbreak severity (Table 4). Note that one explanation for Fool Creek Watershed having relatively low levels of outbreak severity was that 50% of the forested area in the Fool Creek Watershed were harvested between 1954 and 1956 (Troendle and King, 1985). This trend of lower levels of mortality in the cuts was further supported by the fact that most cuts were visible on the map as a very low level of outbreak severity surrounded by higher levels of outbreak severity (Figure 1).

Table 4. A comparison of outbreak severity between watersheds and comparable harvested areas. The East Saint Louis Creek watershed and Lexen Creek watershed serve as controls for the Fool Creek watershed and Deadhorse Creek watershed, respectively. Outbreak severity is measured as a ratio of dead lodgepole pine basal area to total basal area.

Watershed	Mean Outbreak Severity of Watershed	Mean Outbreak Severity of Cuts in Watershed
Fool Creek Watershed	0.10	0.09
East Saint Louis Creek Watershed	0.20	No cuts
Deadhorse Creek Watershed	0.26	0.15
Lexen Creek Watershed	0.25	No Cuts

Elevation explained the most variation in the outbreak severity at each cut (Figure 2). I predicted that “time since cut” would be the most important predictor variable, but that does not

seem to be the case. The regression tree as a whole has a pseudo-R^2 of 0.70. The first split slices harvests into lower elevation and higher elevation cuts. The cuts located at lower elevation experienced higher levels of mortality compared to higher elevation cuts. There is a range of elevation from 2,940 m and 3,012 m where past harvests experienced the highest levels of mortality. Time-since-cut formed the split at the node containing higher elevation cuts. Interestingly, young cuts had a higher mortality (0.15) than older cuts (0.05).

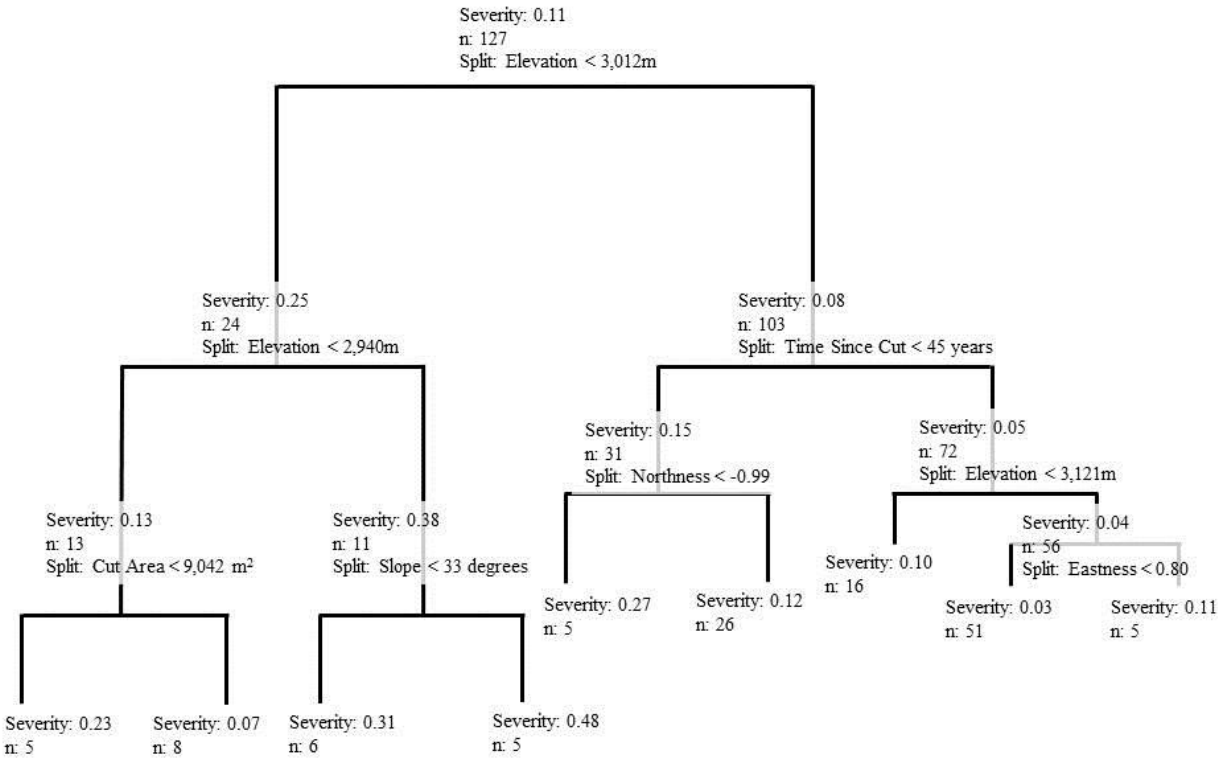

Figure 2. Regression tree showing the relationships between outbreak severity in past harvests relative to topographic attributes and cut stand characteristics. The text at each node tells the average outbreak severity of the cuts at each node, the number (n) of cuts at each node, and the predictor variable split separating the next lower pair of nodes. The height of a branch is proportional to the strength of each split.

4. DISCUSSION

4.1 This Methodology is Effective

I hypothesized that outbreak severity could be accurately modeled at Fraser Experimental Forest with a cross-validation correlation of at least 0.70 and a correlation between predicted and observed outbreak severities at test plots of 0.70 or greater. I found support for my hypothesis when the same data used to train the model were also used to test it using a cross-validation correlation, but I rejected my hypothesis based on the models performance with independent test data. The model performed respectably well and this methodology should be pursued to further improve predictive power. Another attempt to use Landsat TM and ETM+ imagery to map outbreak severity in the same region as Fraser Experimental Forest achieved an R^2 of 0.77 (Meddens and Hicke, 2014). Their methodology used similar predictor variables to those used in my model, but they used a Landsat scene from every year. My methodology is not as accurate as their model, but uses fewer images and may be a less time-intensive option and a better choice in areas where availability of satellite imagery is limited by cloud cover.

The mean texture of ETM+ band 5 calculated on a 3x3 moving window differenced between 2001 and 2010 and the NDMI difference between 2001 and 2010 were the most influential predictor variables (Table 2). While NDMI has been used extensively in mapping forest disturbance from insects (Goodwin et al., 2008; Coops et al., 2010), texture has not been used to map mountain pine beetle outbreaks, so future modeling efforts should consider including image texture to improve model performance. Mean texture may have performed strongly because of the spatially autocorrelated nature of a mountain pine beetle outbreak. If an outbreak is occurring nearby, a pixel is more likely to also experience mortality. Most of the

variability in outbreak severity could be explained by the mean texture difference and NDMI difference, and the addition of more predictor variable only improved predictive power marginally. A quick model of outbreak severity could potentially only include these two predictor variables. However, when the analysis is scaled-up to a larger scale, such as an entire Landsat scene, other predictor variables may become more important.

The best model utilized image differences between 2001 and 2010, ETM+ band 5 from 2010, and elevation. Not one of the image differences or spectral bands from 2012 were used in the final model. This was surprising considering that the plot data were representative of forest conditions in 2012 and 2013. The 2010 imagery may perform better because it was near the tail end of the peak of the outbreak. So, damage from the previous peak years was still fresh. By 2012, the forest response may have obscured outbreak damage. Efforts to map multi-year disturbances, such as insect outbreaks, should consider using imagery from the tail-end of the disturbance to capture the most accurate extent and severity of forest damage. The gain from detecting damage during the waning years of the disturbance may be outweighed by the forest recovery beginning to mask damage from the peak years.

4.2 Harvested Stands were Less Susceptible to Outbreaks

My results supported my hypothesis that forests managed since the 1950s had lower levels of mortality than stands in neighboring watersheds (Table 4). This pattern was also clear from a visual inspection of the mortality map (Figure 1). This conclusion supports the commonly held belief that harvesting can reduce a stand's vulnerability to mountain pine beetle outbreaks by creating stand conditions less favorable for mountain pine beetles (Whitehead et al., 2003).

Studies such as this recording how previously harvested stands faired through the recent outbreak are valuable for policy making and management actions (Six et al., 2014).

Just because harvested areas have lower levels of lodgepole pine mortality, it does not mean that these stands are necessarily "healthier." For example, one of the Deadhorse Creek watershed cuts does not show lodgepole pine mortality, but the plot converted to a bog after being harvested and the forest has not recovered. These sorts of impacts of harvesting are not always evident in a remote sensing analysis such as this one. The long-term condition of these harvested stands should be tracked. Will the stands in Fraser Experimental Forest that experienced higher levels of lodgepole pine mortality be better adapted to future insect outbreaks and climate change due to the strong selective pressure exerted by the beetles (Six et al., 2014)?

Harvested stands are more likely to remain dominated by lodgepole pine, while unharvested areas that were hit by the mountain pine beetle are more likely to have a higher density of subalpine fir and Engelmann spruce advanced regeneration that are released to become the dominant species (Collins et al., 2011). These uncut stands may be more resistant to future mountain pine beetle outbreaks by creating heterogeneous forests. If we do not actively manage previously harvested areas, we may be setting the stage for another large outbreak by perpetuating the presence of even-aged lodgepole pine forests (Whitehead et al., 2003).

4.3 Outbreak Severity Trends

I hypothesized that time since harvest would be the most important variable in explaining outbreak severity and that the more recently harvested the forest, the lower the lodgepole pine mortality would be. I rejected both of these hypotheses. Elevation was the most important

predictor variable (Figure 2) and there is a significant negative Spearman correlation of -0.45 ($p<0.001$) between time since harvest and outbreak severity.

There are a few potential explanations for these trends. The recent outbreak was an order of magnitude larger than previous outbreaks (Six et al., 2014). This intense pressure from such a large beetle population and changing climate lead to previously unexpected mountain pine beetle behavior. Lodgepole pine stands <60 years old were not expected to be infested prior to this outbreak (Shore and Safranyik, 1992). This rule of thumb was refuted by the observation that some of the harvested stands experienced high levels of mortality during this outbreak. The fact that elevation is the primary driver of the regression tree model could mean that temperature, precipitation, and species composition drove the mountain pine beetle population, regardless of tree size. Basically, if there was lodgepole pine present, the beetles infested them. Other studies have found that mortality rates early in the infestation were predictable based on the location of warm, dry sites, and abundant large trees. But, as the outbreak intensified, mortality was associated with species composition (Nelson et al., 2013). Elevation could be acting as a proxy for temperature, precipitation, and species composition in the regression tree (Figure 2). If harvesting is to be used as a mechanism to limit mountain pine beetle outbreak mortality, harvesting should be concentrated on elevations that are most susceptible to mountain pine beetle outbreaks.

It was surprising that of the higher elevation cuts after the first split, the younger cuts experienced higher levels of outbreak severity (Figure 2). This could possibly be explained by a higher density of lodgepole pine in young stands. But, as these higher elevation stands age they could be transitioning to spruce and fir forests that are not vulnerable to mountain pine beetle outbreaks. The split in the lowest left corner of the regression tree is based on the size of the cut.

Smaller stands experienced higher levels of mortality (0.23) compared to larger stands (0.07). This can be explained by the fact that smaller stands can be easily penetrated by beetles from neighboring unharvested stands.

4.4 Caveats

Results from my work should be interpreted with the following caveats in mind. The delineation of past harvests was based on a variety of imperfect sources, so errors in the previously harvested polygons are possible. The type of cut was not considered in the regression tree analysis because this information was not known for many harvests. It would be helpful to know if each harvest was a clearcut or a thinning and if seeding or replanting was done after harvest.

The quality of inference would benefit from field data from the harvested stands to determine species composition, other stand characteristics, and to validate mortality levels. The analysis of outbreak severity in previously cut stands was based on the outbreak severity map which has inaccuracies of its own, so it is possible errors propagated through both analyses. The field data used to test the model were spatially clustered in four watersheds. So, these plots are testing the performance of the outbreak severity model in these watersheds, not across the entire forest.

5. CONCLUSION

Outbreak severity from the recent mountain pine beetle epidemic was modeled at Fraser Experimental Forest using a BRT model with field, GIS, and Landsat 7 ETM+ data. The change in mean texture of ETM+ band 5 and NDMI between 2001 and 2010 proved to be the best predictors of outbreak severity. This work is novel in that it captured outbreak severity for each pixel of a Landsat scene using only two images: one from before the outbreak and one from the declining years of the outbreak.

This map was then used to calculate outbreak severity within cut stands harvested between 1954 and 1985. Stands harvested during this time experienced lower levels of mortality than comparable stands in Fraser Experimental Forest, although there was a large range in outbreak severities between cuts. Elevation of each harvest was an important factor in determining the severity of the outbreak at that stand.

The analysis of outbreak severity in past harvests is just one example of an analysis that is possible with a map of cumulative outbreak severity. Other applications include the study of forest regeneration (Diskin et al., 2011), hydrology (Raffa et al., 2008), fire (Simard et al., 2011), timber harvest (White et al., 2005), and carbon accounting (Pfeifer et al., 2011). This map has the potential to support long-standing research at Fraser Experimental Forest about the impacts of forest disturbances on streamflow. More information about the harvesting methods and stand conditions in the cut stands is necessary before management decisions should be based on this work.

REFERENCES

Alexander, R.A., Troendle, C.A., Kaufmann, M.R., Shepperd, W.D., Crouch, G.L., and Watkins, R.K. 1985. The Fraser Experimental Forest, Colorado: Research Program and Published Research 1937-1985. Rocky Mountain Forest and Range Experiment Station, General Technical Report RM-118. Fort Collins, Colorado.

Bentz, B. J., and Endreson, D. 2003. Evaluating satellite imagery for estimating mountain pine beetle-caused lodgepole pine mortality: current status. In Mountain Pine Beetle Symposium: Challenges and Solutions. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Kelowna, British Columbia, 154-163.

Collins, B. J., Rhoades, C. C., Hubbard, R. M., and Battaglia, M. A. 2011. Tree regeneration and future stand development after bark beetle infestation and harvesting in Colorado lodgepole pine stands. Forest Ecology and Management, 261(11), 2168-2175.

Colorado State Forest Service. 2013. 2013 Colorado Forest Insect and Disease Update: A Supplement to the 2013 Report on the Health of Colorado's Forest. Retrieved from http://csfs.colostate.edu/pdfs/2013FHR-InsectDiseaseUpdate.pdf (Accessed on May 2, 2014).

Coops, N. C., Gillanders, S. N., Wulder, M. A., Gergel, S. E., Nelson, T., and Goodwin, N. R. 2010. Assessing changes in forest fragmentation following infestation using time series Landsat imagery. Forest Ecology and Management, 259(12), 2355-2365.

Crist, E. P. and Cicone, R. C. 1984. A physically-based transformation of Thematic Mapper data---The TM Tasseled Cap. Geoscience and Remote Sensing, IEEE Transactions on, (3), 256-263.

Dennison, P. E., Brunelle, A. R., and Carter, V. A. 2010. Assessing canopy mortality during a mountain pine beetle outbreak using GeoEye-1 high spatial resolution satellite data. Remote Sensing of Environment, 114(11), 2431-2435.

Diskin, M., Rocca, M. E., Nelson, K. N., Aoki, C. F., and Romme, W. H. 2011. Forest developmental trajectories in mountain pine beetle disturbed forests of Rocky Mountain National Park, Colorado. Canadian Journal of Forest Research, 41(4), 782-792.

De'ath, G. and Fabricius, K. E. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology, 81(11), 3178-3192.

Elith, J., Leathwick, J. R., and Hastie, T. 2008. A working guide to boosted regression trees. Journal of Animal Ecology, 77(4), 802-813.

Environmental Systems Research Institute (ESRI), 2012. ArcGIS Desktop: Release 10.1. Redlands, CA.

Exelis Visual Information Solutions, 2014. ENVI 5.1. Boulder, CO.

Gesch, D.B. 2007. The National Elevation Dataset, in Maune, D., ed., Digital Elevation Model Technologies and Applications: The DEM User's Manual, 2nd Edition: Bethesda, Maryland, American Society for Photogrammetry and Remote Sensing, p. 99-118.

Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., and Tyler, D. 2002. The National Elevation Dataset: Photogrammetric Engineering and Remote Sensing, 68(1), 5-11.

Goodwin, N. R., Coops, N. C., Wulder, M. A., Gillanders, S., Schroeder, T. A., and Nelson, T. 2008. Estimation of insect infestation dynamics using a temporal sequence of Landsat data. Remote Sensing of Environment, 112(9), 3680-3689.

Harrell, F. E., Lee, K.L., and Mark, D.B. 1996. Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine 15, 361-387.

Hicke, J. A., and Jenkins, J. C. 2008. Mapping lodgepole pine stand structure susceptibility to mountain pine beetle attack across the western United States. Forest Ecology and Management, 255(5), 1536-1547.

Huang, C., Wylie, B., Yang, L., Homer, C., and Zylstra, G. 2002. Derivation of a tasselled cap transformation based on Landsat 7 at-satellite reflectance. International Journal of Remote Sensing, 23(8), 1741-1748.

Kumar, S., Stohlgren, T. J., and Chong, G. W. 2006. Spatial heterogeneity influences native and nonnative plant species richness. Ecology, 87(12), 3186-3199.

Lu, D., Mausel, P., Brondizio, E., and Moran, E. 2004. Change detection techniques. International Journal of Remote Sensing, 25(12), 2365-2401.

Meddens, A. J., and Hicke, J. A. 2014. Spatial and temporal patterns of Landsat-based detection of tree mortality caused by a mountain pine beetle outbreak in Colorado, USA. Forest Ecology and Management, 322, 78-88.

Morrison, P.H., Smith IV, I.M, and Snetsinger, S.D. 2003. The Natural Communities and Ecological Condition of the Sonoran Desert National Monument and Adjacent Areas. Pacific Biodiversity Institute, Winthrop, Washington, 113 + xvi p. + Vol. 2 (appendices) 395 p.

Morisette, J.T., Jarnevich, C.S., Holcombe, T.R., Talbert, C.B., Ignizio, D., Talbert, M.K., Silva, C., Koop, D., Swanson, A. and Young, N.E. 2013. VisTrails SAHM: visualization and workflow management for species habitat modeling. Ecography, 36, 129-135.

Nelson, K.N., Rocca, M.E., Diskin, M., Aoki, C.R., and Romme, W.H. 2013. Predictors of bark beetle activity and scale-dependent spatial heterogeneity change during the course of an outbreak in a subalpine forest. Landscape Ecology, in press.

Pfeifer, E. M., Hicke, J. A., and Meddens, A. J. 2011. Observations and modeling of aboveground tree carbon stocks and fluxes following a bark beetle outbreak in the western United States. Global Change Biology, 17(1), 339-350.

R Development Core Team, 2013. R: a language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria.

Raffa, K. F., Aukema, B. H., Bentz, B. J., Carroll, A. L., Hicke, J. A., Turner, M. G., and Romme, W. H. 2008. Cross-scale drivers of natural disturbances prone to anthropogenic amplification: the dynamics of bark beetle eruptions. Bioscience, 58(6), 501-517.

Safranyik, L. 2003. Mountain Pine Beetle Epidemiology in Lodgepole Pine. In Mountain Pine Beetle Symposium: Challenges and Solutions. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Kelowna, British Columbia, 33-40.

Shore, T. L., and Safranyik, L. 1992. Susceptibility and risk rating systems for the mountain pine beetle in lodgepole pine stands. Forestry Canada, Information Report 336. Victoria, British Columbia.

Simard, M., Romme, W. H., Griffin, J. M., and Turner, M. G. 2011. Do mountain pine beetle outbreaks change the probability of active crown fire in lodgepole pine forests? Ecological Monographs, 81(1), 3-24.

Singh, M., Malhi, Y., and Bhagwat, S. 2014. Evaluating land use and aboveground biomass dynamics in an oil palm–dominated landscape in Borneo using optical remote sensing. Journal of Applied Remote Sensing, 8(1), 083695.

Six, D. L., Biber, E., and Long, E. 2014. Management for Mountain Pine Beetle Outbreak Suppression: Does Relevant Science Support Current Policy?. Forests, 5(1), 103-133.

Skakun, R. S., Wulder, M. A., and Franklin, S. E. 2003. Sensitivity of the thematic mapper enhanced wetness difference index to detect mountain pine beetle red-attack damage. Remote Sensing of Environment, 86(4), 433-443.

Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P., and Macomber, S. A. 2001. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? Remote sensing of Environment, 75(2), 230-244.

Troendle, C.A. and King, R.M. 1985. The Effect of Timber Harvest on the Fool Creek Watershed, 30 Years Later. Water Resources Research, 21(12), 1915-1922.

USDA Forest Service, 2013. Forest Insect and Disease Aerial Survey Data, Rocky Mountain Region. Retrieved from http://www.fs.fed.us/r2/resources/fhm/aerialsurvey/download/ (accessed October 1, 2013).

Walker, W. S., Kellndorfer, J. M., LaPoint, E., Hoppus, M., and Westfall, J. 2007. An empirical InSAR-optical fusion approach to mapping vegetation canopy height. Remote Sensing of Environment, 109(4), 482-499.

Walter, J. A. and Platt, R. V. 2013. Multi-temporal analysis reveals that predictors of mountain pine beetle infestation change during outbreak cycles. Forest Ecology and Management, 302, 308-318.

White, J. C., Wulder, M. A., Brooks, D., Reich, R., and Wheate, R. D. 2005. Detection of red attack stage mountain pine beetle infestation with high spatial resolution satellite imagery. Remote Sensing of Environment, 96(3), 340-351.

Whitehead, R. J., Safranyik, L., Russo, G., Shore, T. L., and Carroll, A. L. 2003. Silviculture to reduce landscape and stand susceptibility to the mountain pine beetle. In Mountain Pine Beetle Symposium: Challenges and Solutions. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Kelowna, British Columbia, 233-244.

Wulder, M. A., Dymond, C. C., White, J. C., Leckie, D. G., and Carroll, A. L. 2006a. Surveying mountain pine beetle damage of forests: A review of remote sensing opportunities. Forest Ecology and Management, 221(1), 27-41.

Wulder, M. A., White, J. C., Bentz, B., Alvarez, M. F., and Coops, N. C. 2006b. Estimating the probability of mountain pine beetle red-attack damage. Remote Sensing of Environment, 101(2), 150-166.

APPENDIX

R Code

```
# Boosted Regression Trees in R

################ Note to reader ####################

# Modified by Nick Young 09/14/2011

# Customized by Tony Vorster 06/20/2014

# This is a cut-down version of a tutorial prepared by Jane Elith and John Leathwick,

# to show how to use it to fit boosted regression tree models.

# There is a word document provided by Jane Elith and John Leathwick that may provide more

# explanation

################Load and check data###############

# Clear memory from R

rm(list=ls())

# Get Source code

source("I:\\Thesis\\Model\\BRT\\brt.functions.R")

# Data contain the response of all the species collected and the predictor variables

BRT.data=read.csv("I:\\Thesis\\Model\\BRT\\BRT_BA_Jun19.csv")

#Data check command to see if data appear correct

head(BRT.data) #This gives the first 6 rows of the data

names(BRT.data) #This lists all the column names of the data

################Organize data###################

# Look at the environmental variables we are going to use for the model
```

```
predictor_col<-c(10,12,14,15,18,20)
names(BRT.data[predictor_col])
var.bind<-BRT.data[predictor_col]
# Create a list of all variable names to be used later
var.names<-names(var.bind)
# Set response to model
response<-3
Y=BRT.data[,response]
# Combine response with predictors to go into gbm
# The numbers index which columns to include
data.gbm<-cbind(Y,var.bind)
################Find Model Settings#######################
# Run gmb.step to identify what learning rate and how many trees to use
# The output will identify how many trees to include in the model or if you need to change
# the learning rate or number of trees for a better model
# Call to load the gbm package
library(gbm)
# You made need to change the learning.rate and tree.complexity if the results say you should
fit.step<- gbm.step(
  data=data.gbm,
  gbm.x = 2:7, #these are the columns that contatin the predictors
  gbm.y = 1,  #your response
  family = "gaussian",
```

```
  tree.complexity = 3,
  learning.rate = 0.0025,
  bag.fraction = 0.5
)
############Finalize the model#################
# This function identifys how many and which vairables to remove from the model
fit.simp <- gbm.simplify(fit.step)
# This creates the final model. Notice that you can set how many variables to
# drop from the gbm.simplify function to specificy what variables to include
fit.step.simp <- gbm.step(data.gbm, gbm.x = fit.simp$pred.list[[1]], gbm.y = 1, family =
"gaussian", tree.complexity = 3, learning.rate = 0.0025, bag.fraction=0.5)
###########Examine the Output#################
# Graph the response variables
par(mfrow=c(3,4))
gbm.plot(fit.step.simp, n.plots=12, write.title = F)
gbm.plot.fits(fit.step.simp)
find.int <- gbm.interactions(fit.step.simp)
# Calculate relative influence
summary(fit.step.simp)
########################TEST DATA#############################
#This is where you need to bring in your test data
data.test<-read.csv("I:\\Thesis\\Model\\BRT\\BRT_SnowCourse_Jun19.csv")
Y.test<-data.test[,response]
```

```
head(Y.test)
# Make predictions for the test data
pred.test <- predict.gbm(fit.step.simp, data.test, fit.step.simp$n.trees)
pred.observ <-cbind(Y.test,pred.test)
# Calculate correlation between observed vs. predicted values from BRT
# Pearson correlation
correlation.p<-cor(Y.test,pred.test)
# Spearman correlation
correlation.s<-cor(Y.test,pred.test,method="spearman")
correlation.p
correlation.s
# Run linear regression between observed and predicted to calculate R2
reg.predct.observe<-lm(Y.test~pred.test)
summary(reg.predct.observe)
# Plot observations and model predictions
Xval<-Y.test
Yval<-pred.test
plot(Xval,Yval, xlab="Observered abundance",ylab="Predicted abundance")
#####Create model prediction ASCII####
# Open Source R code for processing ascii by line
source("I:\\Thesis\\Model\\BRT\\process_asc_by_line_v3.r")
myfun<-function(x)round(unlist(predict.gbm(fit.step.simp, x, fit.step.simp$n.tree)),4)
# Predictor and ASC names (the ASCIIs and predictor names must match)
```

```
# Check names
row.names(fit.step.simp$contributions)<-fit.step.simp$contributions[,1]
Thenamestable<-fit.step.simp$contributions
var.names<-row.names(Thenamestable)
var.names
# Path to a folder containing the ASCII predictor files
fpath <-"I:\\Thesis\\Model\\SAHM\\Output\\June17\\ascii\\"
# Spatial prediction function. Change output name each run.
proc.asc.byline(var.names,fpath,myfun,n=100,outfile="BRT_Jun20.asc")
#################Regression Tree in R###################
CART.data=read.csv("I:\\Thesis\\Model\\CART\\Cut_Attributes.csv")
attach(CART.data)
library(tree)
covariates <- CART.data[c(4,6,19,20,22,23)]
cor(covariates)
# Regression tree model
rtree1 <- tree(MEAN_Severity~YearsSinceCut+Cut_Area + Mean_Elev + Mean_Slope +
Northness + Eastness,data=CART.data)
summary(rtree1)
plot(rtree1)
# Run cross-validation to select best tree size, i.e. prune the tree
# 10 k-fold cross validations
rtree1.cv <- cv.tree(rtree1, K=10)
```

```
plot(rtree1.cv)

snip.tree(rtree1)

# You can visually inspect plot and select size based on min deviance or use code below

opt.trees = which(rtree1.cv$dev == min(rtree1.cv$dev))

best.leaves = min(rtree1.cv$size[opt.trees])

rtree1.p <- prune.tree(rtree1,best=best.leaves)

plot(rtree1.p)

text(rtree1.p,splits=TRUE, label="yval", all=FALSE)

summary(rtree1.p)
```